

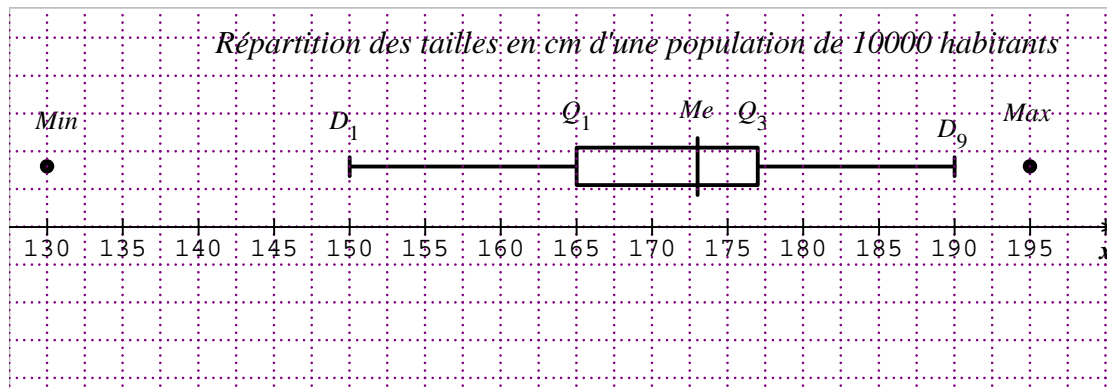
I. Rappels de Statistiques

Les séries statistiques sont généralement composées d'un grand nombre de valeurs. Il convient donc de résumer une série à l'aide de quelques indicateurs privilégiés : moyenne, écart-type, médiane, quartile...

Lorsque la série comporte un très grand nombre de valeurs, médiane et quartile sont des indicateurs insuffisants, et on définit alors les *déciles* ou les *centiles*.

Par exemple, les déciles sont obtenus en partageant la population en 10 sous-populations de même effectif : le neuvième décile D_9 est par exemple le nombre tel que 90% de la population est inférieur à D_9 et donc 10% de la population lui est supérieur.

Dans la majorité des exercices de cette année, ces indicateurs de dispersion seront représentés à l'aide d'un diagramme en moustache.



Exemple de lecture du diagramme à moustache :

- Les tailles minimales et maximales observées dans la population sont respectivement 130cm et 195cm.
- 10% de la population mesure moins de 150 cm, 10 % mesure plus de 190cm.
- 25% de la population mesure moins de 165 cm, 25 % mesure plus de 177cm.
- 50% mesure moins de 173 cm.
- ...

Remarque : Vu que seulement 10% des valeurs de la série sont au dessus de D_9 , on dira aussi que ces valeurs sont *marginales* (en quelque sorte **exceptionnelles**).

Notons que les déciles seront, dans nos exercices des données de l'énoncé, préalablement estimés par ordinateur.

II. Test d'adéquation à une loi équirépartie

Adéquation est à rapprocher des mots adéquat, approprié, convenable ou juste.

Loi Equirépartie : une loi *équirépartie* est une loi dans laquelle chaque événement élémentaire à la même probabilité de réalisation. Ils sont *répartis* de manière *équitable*.

Par exemple, si X désigne la variable aléatoire égale à la face d'un dé équilibré, X suit une loi équirépartie : chaque événement élémentaire à la même probabilité ($\frac{1}{6}$) de sortie.

Un test d'adéquation sera un test qui nous permettra de vérifier, à partir des résultats d'une expérience E, s'il est juste ou convenable de considérer qu'un phénomène donné suit une loi équirépartie (avec un certain risque d'erreur).

Peut-on supposer de manière raisonnable, à partir de 50 lancers réalisés, que notre dé (ou pièce) est équilibré ?

A partir de 2000 naissances réparties par trimestre parues lors d'une enquête, peut-on supposer de manière raisonnable que l'on naît avec la même probabilité quelque soit le trimestre de l'année ?

Les accidents mortels dus à l'alcoolisme sont-ils uniformément répartis dans la semaine ?

III. Méthode de test

Nous allons présenter la méthode à travers l'étude d'un exemple.

- Question : On a une pièce, on veut savoir si elle est équilibrée.

- Expérience E : On jette (donc) 50 fois cette pièce. On observe « 30 Piles et 20 Faces ».

- Calcul de la « distance entre l'expérience et la théorie » :

→ Expérience : Sur 50 lancés, la fréquence de l'évènement « 30 Piles » est $f_p = \frac{30}{50} = \frac{3}{5}$ et celui de « 20 Faces » est

$$f_f = \frac{20}{50} = \frac{2}{5}.$$

→ Théorie : en théorie, pour une pièce équilibrée, la fréquence de ces deux évènements est de $f_p = f_f = \frac{1}{2}$.

→ Comme pour un calcul de distance classique, la distance « Réalité - Théorie » est donc de

$$D_{obs}^2 = \sum_{issues\ element} (f_{obs} - f_T)^2 = \left(\frac{3}{5} - \frac{1}{2}\right)^2 + \left(\frac{2}{5} - \frac{1}{2}\right)^2 = \frac{1}{50}.$$

Pour travailler avec des entiers, on calcule par exemple $5000D_{obs}^2 = 100$.

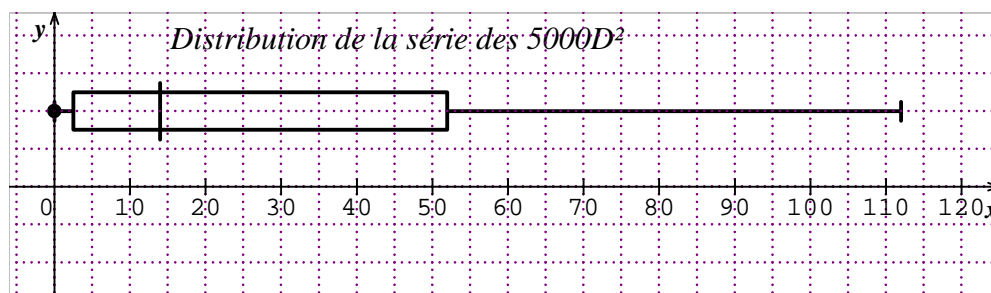
- Simulation : on veut maintenant savoir si cette valeur de $5000D_{obs}^2$ est cohérente avec ce qu'aurait donnée une pièce vraiment équilibrée dans la pratique.

Est-elle marginale (rare) ou une pièce équilibrée aurait donné une valeur semblable ?

A l'aide d'un ordinateur, on simule donc 50 jets d'une pièce équilibrée : on calcule alors $5000D^2$.

On recommence cela un très grand nombre de fois, par exemple 10 000 fois.

On dispose donc de 10 000 valeurs de $5000D^2$: on détermine alors le neuvième décile de cette série.



- On lit $D_9 \approx 112$.

Cela signifie que pour une pièce équilibrée, seulement 10% des valeurs sont au dessus de 112 (donc marginales).

On peut aussi dire que dans 90% des expériences réalisées avec une pièce équilibrée, $5000D^2 < 112$.

Vu que, suite à notre expérience, $5000D_{obs}^2 = 100 < 112$, on peut affirmer, avec une marge d'erreur de 10%, que notre une pièce équilibrée.

Remarque : contrairement à ce qu'on pourrait penser, si on travaille avec le 99^{ème} centile, cad avec une marge d'erreur de 1%, la plupart des pièces testées seront équilibrées ! En effet, avec une telle marge, seuls 1% des cas seront considérés comme marginaux, ce qui rallonge de manière significative la zone des 90% des résultats acceptés...

On peut aussi préciser que plus D^2 est petit, plus il semble légitime de supposer que l'expérience suit une loi équirépartie (puisque D^2 mesure la distance « théorie - expérience »).

IV. Exercices de Bac

Antilles Guyane, Septembre 2004, QCM question 3.

France, Juin 2006.

Allez consulter le site !